

**INM431 Machine Learning
Coursework
Submission deadline: Sun, 29 Nov 2015, 5:00PM**

Introduction

On completing this coursework you should be able to:

1. Specify a Machine Learning (ML) solution to a data analysis problem;
2. Adjust the ML model parameters and explain how ML seeks to solve the problem;
3. Apply, compare and contrast, and critically evaluate two ML models.

The subject of the above learning outcomes follows lectures 1 to 8 and associated materials and exercises.

Information

This coursework is worth 30% of the module mark.

Submission

Only submissions through Moodle will be accepted. Please submit all file(s) needed to run your experiments (as a single zip file) and a poster in A1 format (as a pdf file). Either submit the dataset as part of the zip file or add an http link to the poster stating where the dataset can be downloaded from.

No late submissions will be accepted. You are strongly advised to submit draft versions of the coursework well before the deadline. Don't leave the submission to the last minute!

Teamwork

This coursework is to be done in groups of 2 students. One student should submit all the coursework files (zip and pdf) and a self/peer-assessment txt file. The other student should submit a self/peer-assessment txt file only. The txt file should have the format:

<Student_name_1> : <evaluation_score>
<Student_name_2> : <evaluation_score>

where evaluation_score in {1,2,3,4} means:

- 1 = didn't contribute
- 2 = contributed somewhat
- 3 = contributed
- 4 = outstanding contribution

Marking scheme

Solutions will be marked according to the following criteria. Code: syntactic correctness (5%), organization and clarity of comments (10%), appropriate use and sophistication of methods (10%). Poster: brief description and motivation of the problem (5%), initial analysis of the data set including basic statistics (10%), brief summary of the two ML models with their pros and cons (10%), hypothesis statement (5%), description of choice of training and evaluation methodology (5%), choice of parameters and experimental results (10%), analysis and critical evaluation of results (25%), lessons learned and future work (5%). Marks will be adjusted according to the self and peer-assessments. You are not being marked on how good the results are, neither on your poster presentation skills. What matters is that you try something sensible and clearly present the problem, your method, what you did, and what the results were. The

poster session will be an opportunity for you to show us all the work you did. Posters will be marked based on their clarity, relevance to the problem statement, relationship to previous work and references, presentation and comparative analysis of experimental results with the use of readable and informative figures and tables.

General Guidelines

The idea with this coursework is to give you some experience carrying out and presenting a piece of research in machine learning. What we expect to see is a task that you describe clearly, relate to existing work, implement and test on a dataset using two ML models. To do this you will need to write or change code, run it on some data, make some figures, read a few background papers, collect some references, compare and contrast results, and create a poster describing your task, the algorithm(s) you used and the results you've obtained. As a rough rule of thumb, spend about a week's worth of work (spread out over a longer time to allow the computers to do some work in the interim!), and about three days preparing the poster and rehearsing your presentation. We encourage you to work in pairs.

Specific Requirements

You are asked to compare two machine learning algorithms in practice, when applied to some data. You may also propose a new algorithm, in which case you still should compare it to one other approach. Select the data and algorithm(s) wisely! Your poster should include at least two figures which graphically illustrate quantitative aspects of your results, such as training/testing error curves, learned parameters, algorithm outputs, etc. It should also include at least 5 references to research papers or book chapters.

Coursework Proposal

Feel free to bring a brief coursework proposal (1-2 paragraphs) to class for discussion at the tutorial sessions. Your proposal can state, e.g. the problem that you are planning to solve, with associated dataset(s), and list any software you will need to write, use or change, or a couple of papers you're planning to read that will help you complete the coursework.

Extenuating Circumstances

If you are not able to submit your coursework for medical reasons or any serious personal reasons beyond your control you should contact the programmes office and fill an extenuating circumstances form. Medical certificates will be required.

Plagiarism

If you copy the work of others (either that of another student or of a third party), with or without their permission, you will score zero marks and further disciplinary action may be taken.